

■ Neurocognitive Development and Predictors of L1 and L2 Literacy Skills in Dyslexia: A Longitudinal Study of Children 5–11 Years Old

Turid Helland* and Frøydis Morken

Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway

The aim of this study was to find valid neurocognitive precursors of literacy development in first language (L1, Norwegian) and second language (L2, English) in a group of children during their Pre-literacy, Emergent Literacy and Literacy stages, by comparing children with dyslexia and a typical group. Children who were 5 years old at project start were followed until the age of 11, when dyslexia was identified and data could be analysed in retrospect.

The children's neurocognitive pattern changed both by literacy stage and domain. Visuo-spatial recall and RAN appeared as early precursors of L1 literacy, while phonological awareness appeared as early precursor of L2 English. Verbal long term memory was associated with both L1 and L2 skills in the Literacy stage. Significant group differences seen in the Pre-literacy and Emergent literacy stages decreased in the Literacy stage.

The developmental variations by stage and domain may explain some of the inconsistencies seen in dyslexia research. Early identification and training are essential to avoid academic failure, and our data show that visuo-spatial memory and RAN could be suitable early markers in transparent orthographies like Norwegian. Phonological awareness was here seen as an early precursor of L2 English, but not of L1 Norwegian. © 2015 The Authors. *Dyslexia* published by John Wiley & Sons Ltd.

Keywords: visuo-spatial skills; rapid naming; phonological awareness; short term memory; working memory; long term memory

INTRODUCTION

Dyslexia is defined as a disorder that affects literacy and language skills, it is present at birth, it is life long, it is characterized by a number of neurocognitive difficulties and it is resistant to conventional training (The British Dyslexia Association, 2007). Also, it is seen in all ranges of intellectual capacities, and it is characterized by multiple domains of impairment (Pennington & Bishop, 2009). The way it expresses itself changes with age, it can be understood at different levels and it is highly idiosyncratic.

According to Frith, dyslexia can be analysed at a biological, a cognitive and a symptomatic level, which are all influenced by the environmental level (Frith, 1995). Frith characterizes the cognitive level as the bridge between the

*Correspondence to: Turid Helland, Department of Biological and Medical Psychology, University of Bergen, Jonas Lies vei 91, N-5009 Bergen, Norway. E-mail: turid.helland@psybpb.uib.no

symptomatic and the biological levels, and defines dyslexia as an impairment of neurocognitive origin (Frith, 1999). A multitude of neurocognitive benchmarks of dyslexia are described in the literature, and studies report of impairments within phonological awareness (PA) (see e.g. Melby-Lervåg & Lervåg, 2011; Vellutino, Fletcher, Snowling, & Scanlon, 2004), rapid automatized naming (RAN) (see e.g. Landerl & Willburger, 2010; Norton & Wolf, 2012), short term memory (STM) (see e.g. Beneventi, Tønnesen, & Ersland, 2009; Treacy, Steve, & Martine, 2013), working memory (WM) (see e.g. Baddeley & Hitch, 1974; Helland & Asbjørnsen, 2004; Torgesen & Houck, 1980), visual skills (VS) (see e.g. Heiervang & Hugdahl, 2003; Helland & Asbjørnsen, 2003; Stein & Walsh, 1997; Valdois, Lassus-Sangosse, & Lobier, 2012; Vidyasagar & Pammer, 2010; Wang *et al.*, 2014), executive functions (EF) (see e.g. Gooch, Snowling, & Hulme, 2011; Helland & Asbjørnsen, 2000; Reiter, Tucha, & Lange, 2005), long term memory (LTM) (see e.g. Baddeley, 2000; Kibby & Cohen, 2008) and language comprehension (see e.g. Helland, 2007; Helland & Kaasa, 2005; Snowling, Bishop, & Stothard, 2000; Xiao & Ho, 2013).

For many years the theory of impaired phonological skills has had a dominant position, and has also been seen as a causal explanation of dyslexia (Vellutino *et al.*, 2004). But the theory has also been questioned, mainly because it has been given a definitional status without being clearly defined (see e.g. Lyon, Shaywitz, & Shaywitz, 2003; Ramus & Szenkovits, 2008; Tønnessen, 1997; Uppstad & Tønnessen, 2007). Studies often include STM and RAN in the concept of PA, but it has become more and more clear that these are independent predictors of dyslexia (see e.g. Kirby, Georgiou, Martinussen, & Parrila, 2010; Norton & Wolf, 2012; Ramus, 2003). The increasing research on dyslexia in different orthographies points to a reevaluation of core deficits in dyslexia (see e.g. Blomert & Willems, 2010; Landerl *et al.*, 2013; Torppa, Lyytinen, Erskine, Eklund, & Lyytinen, 2010). Studies of L1 and L2 learning indicate that orthographic transparency or script plays an important role for how PA in literacy acquisition should be evaluated (see e.g. Chung & Ho, 2010; Helland, 2008; Helland & Kaasa, 2005; Miller-Guron & Lundberg, 2000).

Obviously, the significance of each of the dyslexia 'benchmarks' is inconclusive. However, they are often analysed within the Model of Working Memory, with its four main components, the phonological loop, the visuo-spatial sketchpad, the central executive and the episodic buffer (Baddeley, 2003). The function of the phonological loop is to perceive and encode auditory and visual input, to be stored or recorded and sent through the phonological output buffer to verbal output. The role of the visuo-spatial sketchpad is less clear as to language functions, but it has a capacity to perceive, encode, hold and manipulate visuo-spatial representations over a short period of time. The central executive controls and regulates the WM system, while the episodic buffer binds the three components to LTM. These functions are basic for a wide range of neurocognitive activities, and especially for the language system (see e.g. Baddeley, 2003; Gathercole, Alloway, Willis, & Adams, 2006).

Literacy learning is heavily based on a child's language skills, but it differs from verbal language learning in that visual formats and hence VS skills are required. Letters have to be visually perceived, encoded, stored and retrieved from memory to be used for reading and writing. In preschool assessments of impaired language development, verbal tasks are typically used (see i.e. Bishop & Leonard, 2000; Schwartz, 2009), while visual tasks have been given minor attention. Although

there is little indication that the sketchpad component in WM plays a significant role in language learning, there is a tendency for younger children, in contrast to older children, to remember pictorial information in visual rather than verbal format (Gathercole & Baddeley, 1993). Thus, studies of pre-school children from different cultures indicate that early at-risk factors of dyslexia should include tests of VS (see i.e. Wang *et al.*, 2014).

As pointed out by Halliday (2014) this complex interaction of many factors may explain why different publications on dyslexia yield different and at times contradictory results. A majority of dyslexia studies are cross sectional, and comparatively few studies have a longitudinal design. However, prominent researchers within the field have claimed that well-constructed longitudinal studies are the best way to sort out all the inconsistencies found in dyslexia research (Dehaene, 2009; Goswami, 2003). According to Goswami developmental designs in dyslexia research should be within a neuroconstructivist framework, meaning that the lowest level of impairment should be identified as early as possible, providing a basis for the understanding of the development of higher level cognition (Goswami, 2003, 2008). This is in accordance with the described aims of the present longitudinal study. A problem with longitudinal studies is that they may be hard to compare because of different selection criteria. A genotype criterion, as used in i.e. Torppa *et al.* (2010), is problematic in the sense that not all children of dyslexic parents develop dyslexia, and not all individuals who develop dyslexia have dyslexia in their close biological family. A phenotype criterion, as used in i.e. Shaywitz *et al.* (1999), is problematic in the sense that obviously dyslexia cannot be identified until after reading and writing training in school seem to fail, and thus studies based on this selection criterion miss out on pre-literacy assessment. An endophenotype criterion (see i.e. Moll, Loff, & Snowling, 2013) combines these two criteria and can be used to find early markers of developmental dyslexia.

In line with developmental frameworks as described by i.e. Frith and Ehri, a child's roadway to literacy can be separated into three stages (Frith, 1986; Ehri, 1987). The first one is the Pre-literacy stage ahead of formalized literacy training in school. The child recognizes frequently seen logos, but cannot read the individual letters or synthesize them into words. In their play they often 'draw' letters and 'read' aloud what they have drawn or 'written'. Next is the Emergent literacy stage when the child receives formal literacy training. The child is taught the letters, the grapheme/phoneme correspondence, how to synthesize these elements into words, or, reversely, how to analyse the words by their graphemes. The third stage is the Literacy stage. At this stage reading is automatized, the child reads the whole word or several words at a time, and writing floats easily without analysing the grapheme/phoneme correspondences of each word. At this stage literacy is no longer a skill to be learned in itself, but a tool for academic learning.

In the western world L1 literacy learning is often paralleled with L2 learning. Over the past decades English has increasingly become a lingua franca and a key to international communication. In many countries it is taught from the first grades on as a second language (L2). L2 learning has to rely on the L1 steps into literacy, but cannot be described in stages in the same manner as L1, because the literacy 'code' should preferably be cracked in L1, and differences in linguistic traits, cultures, time and circumstances of learning, school curriculum and environment make L2 literacy learning different from L1 literacy learning. It depends both upon learning in school and learning via mass media and popular culture (Helland &

Kaasa, 2005; Helland, 2008; Vulchanova, Foy, Nilsen, & Sigmundsson, 2014; Geva, Yaghoub Zadeh, & Schuster, 2000). In Norway, English as L2 is introduced already in the 1st grade, but is formalized as part of the general school curriculum from the 2nd grade on. The children are introduced to English by listening to the phonemes of the language and trying to recognize them by graphemes in simple utterances, and by listening to and comprehending rhymes, riddles, songs, fairy tales and short stories. From the 4th to the 7th grades the aim is to improve oral communication skills, to be able to understand the connection between English phonemes and graphemes, and to read and write simple texts (Norwegian Directorate, 2013). The focus has been on oral and general communication skills. Implicit learning strategies have been the focal point—demanding setups of daily situations in the classroom, where the pupil has to listen to and produce language, and try to master the pragmatics of the L2. The use of orthographic images (flashcards) is not established as an L2 teaching method. Criticism has been raised as no formal requirements are put on the teacher's competence in English as L2 in the elementary school (i.e. grades 1–7). As the exposure to English out of the school setting is relatively high in Norway, the mass media play an important role as most programmes and movies are broadcast in the original language with subtitles and no voice-overs. Therefore, it is hard to say what contributes most to the children's English learning—school curriculum or mass media. For further discussion of this topic, please see Dahl and Vulchanova (2014).

The present study reports from the Bergen Longitudinal Dyslexia Study (also called the 'Speak up!' project, see <http://www.uib.no/en/project/speakup>), which followed a group of at-risk children and matched controls from they were 5 to 11 years old, when dyslexia was diagnosed and the collected data could be analysed retrospectively (Helland, Plante, & Hugdahl, 2011). The children were subjected to a wide range of assessments, covering language and literacy development, in both first language (L1, Norwegian) and second language (L2, English) and neurocognitive skills. Brain scanning was performed at ages 6, 8 and 12 (Clark *et al.*, 2014; Morken, Helland, Hugdahl, & Specht, 2014; Specht *et al.*, 2009). Our analyses at the symptomatic and biological levels showed both group differences and changes by age. In the present study the 'bridge' between these two levels (cp. Frith's terminology), the cognitive level, is described through analyses of longitudinal neurocognitive data.

Here we wanted to (1) assess the correspondence between the neurocognitive development at each of the three stages (Pre-literacy, Emergent literacy and Literacy) and age 11 L1 and L2 literacy skills; (2) find valid neurocognitive precursors of L1 and L2 literacy development at each of the three stages and (3) compare the neurocognitive development in the dyslexia group and the typical group across the three stages of L1 and L2 literacy acquisition.

METHOD

Compared to many other countries, socio-economic differences in Norway are minor (Halvorsen & Stjernø, 2008). The municipalities participating in the project consist of small urban or rural communities, with comparable living standards and with minor ethnical diversity. The general pedagogical approach in Norwegian pre-schools is focused on free play and activities as singing, storytelling, drawing and

outdoor playing (Norwegian Directorate, 2005). The Norwegian school system is public, unitary and for all with an overall educational ideology of active student inclusion and individually adapted teaching (Norwegian Directorate, 2010–2011). Norwegian has two very similar official orthographies representing two dialectal varieties; 'Bokmål' and 'Nynorsk'. The children were tested with the orthography they were taught in school. The educational level of the mothers was 13.4 years ($SD = 2.5$) of schooling with no statistical difference between the subsequent groups that were established (see Pancsofar & Vernon-Feagans (2006) for a further evaluation of parents' influence on the language development of their children).

Participants

The study was approved by the Regional Committee for Medical Research Ethics in Western Norway (REK-Vest) and Norwegian Social Science Data Services (NSD). The participants were children from Norwegian pre-schools selected on the criteria that all four counties in western Norway (Rogaland, Hordaland, Sogn og Fjordane and Møre og Romsdal) and both urban and rural districts should be represented. In all nine preschools with a total of 120 five-year-old children were selected by the respective county community authorities. All parents, teachers and clinicians from the preschools were orally informed about the project at local meetings. They were also informed that an at-risk group of developmental dyslexia and a control group formed by fewer than half of the children would be selected to take part in the study. A letter of consent was then sent to the parents and teachers of the 120 children by September in the year when the children were five years old.

Criteria of exclusion at project start were impaired sight or hearing, intellectual disability according to DSM-IV criteria (APA, 1994) and diagnoses of any other impairment included in the DSM-IV (various syndromes such as ADHD, autism spectrum disorders and neurological impairments) as reported by parents. All participants had to have Norwegian as their first language. Four children did not meet the inclusion criteria. All data collection points reported in this study were from the fall terms, October through November. In addition, the children received periodical data-based training in accordance with the national pre-school and school curriculum during their last months of preschool and in 1st and 2nd grades (Helland *et al.*, 2011b), and fMRI sessions at ages 6, 8 and 12 were carried out during spring terms (Clark *et al.*, 2014; Morken, Specht, Hugdahl, & Helland, 2014; Specht *et al.*, 2009).

Participants at each literacy stage

Pre-literacy stage, age 5–6 (pre-school, 1st grade): Informed consents were received from 109 parents (90.8%) who consented to let their children participate in a further selection of at-risk children and matched controls to be followed for a period of four years. A total of 109 Risk Index questionnaires (RI-5) (please refer to Helland *et al.* (2011a) and Helland (2015) for details on RI-5) were sent to and completed by parents and preschool teachers. 105 children met the criteria of inclusion, and they were sorted into a Risk group ($N = 26$; $M = 13$, $F = 13$) and a matched Control group ($N = 26$; $M = 13$, $F = 13$). The remaining children ($N = 53$) were not followed longitudinally (Helland *et al.*, 2011a).

Emergent literacy, ages 7–8 (2nd and 3rd grades): Three children were withdrawn by their parents during this period. Thus, the Risk group counted 25 children; (M = 13, F = 12), and the Control group 24 children (M = 12, F = 12).

Literacy stage, age 11 (6th grade): Informed consents to take part in the follow-up study were received from 42 children and their parents. Thirteen children (M = 5, all from the original at-risk group; F = 8, six from the original at-risk group) were identified with dyslexia. The remaining 29 (M = 17; F = 12) were identified with typical literacy skills, with no difference in literacy scores between the at-risk children who did not develop dyslexia and the original control group. The attained L1 and L2 literacy scores were used as baseline in the present study (see also Helland, Plante, *et al.*, 2011; Morken & Helland, 2013).

Data Collection

The assessments took place every fall when the children were 5 to 8 years old. Then there was a pause, and the children were assessed again when they were 11 years old (6th grade). When the children in this study were in the 5th grade, they took part in the national tests, and results on a community basis were comparable to national means (Norewegian Directorate, 2009).

The children attended nine different preschool when they were 5 years old. After preschool the children entered 11 different schools in the four counties when they were 6 years old. In accordance with school curriculum literacy training started in 2nd grade when they turned 7. The period from 2nd through 3rd grades focused on emergent literacy, and from 4th grade on literacy was expected to be a tool for learning, implying that reading and writing have become automatized skills. Thus, the curriculum is in accordance with the three phases of literacy acquisition described earlier. For the children in the present study, English as L2 was gradually introduced after the first L1 literacy training had taken place.

Missing Data

Because of illness not all participants took part in all assessments during the 7-year span of the project. Thus, the number of participants in the Typical group varies from 29 to 27, with missing data from different subjects at different times.

Assessments

Baseline

At project start at age 5 all children were tested with the Wechsler Preschool and Primary Scale of Intelligence™—Third Edition (WPPSI™—III) (Wechsler, 2002). The retrospective scores of the WPPSI and the RI-5 in the Dyslexia and Typical groups are shown in Table 1.

Literacy

L1 Word reading. Nonword reading and single word reading were assessed by the STAS (Standardisert Test i Avkodning og Staving [Standardized Test of Decoding and Spelling]) (Klinkenberg & Skaar, 2001). For nonword reading a list of 85 words was to be read within a limit of 40 s. For single word reading four lists of 85 words

Table 1. Dyslexia vs typical group, age 5: WPPSI™-III and RI-5

Task	Dys	SD	Typ	SD	t-value	df	p	Cohen's <i>d</i>	N Dys	N Typ
WPPSI, age 5										
5/VIQ	100.3	15.92	103.5	10.57	−0.757	40	.45	−0.23	13	29
5/PIQ	99.7	18.24	102.3	23.11	−0.356	40	.72	−0.12	13	29
RI-5	20.6	14.50	9.6	8.07	3.172	40	.003	0.94	13	29

Notes. Baseline scores: VIQ = verbal scores; PIQ = performance scores from WPPSI; RI-5 (Risk Index questionnaire): high scores indicate risk of developmental dyslexia, low scores indicate no risk (see Helland et al., 2011).

Bold data indicate significance.

were to be read aloud within a time limit of 40 s each. The test is standardized from 2nd to 10th grade. One point was given for correct response, 0 for incorrect response. Maximum score is 85 points for the nonword reading test read within 40 s and 340 points for the single word reading test read within 160 s. For the 6th grade, the norm score is 63 (SD 19) for nonword reading and 205 (SD 55) for word reading.

L1 Text reading. The Carlsten Reading Test, grade 6 (Carlsten, 2002) was used to assess text reading speed and comprehension. The test consists of two short stories written by two different authors in the two official Norwegian orthographies, with 1165 words in the 'Bokmål' version and 1238 in the 'Nynorsk' dialect version. The two versions have 27 and 28 cloze tasks, respectively, in which the reader is asked to mark the correct alternative of three printed words in a bracket. The test has tentative norm scores and is constructed for classroom testing of reading speed and reading errors within a time limit of 10 min. An acceptable reading score at this age is 100 words/min and above. However, because we wanted a more accurate reading and comprehension score, a new reading score was calculated, showing words per minute in comprehended text reading. See Helland et al. (2011a) for a detailed description.

L1 Word spelling was assessed by the STAS (Standardisert Test i Avkodning og Staving [Standardized Test of Decoding and Spelling]) (Klinkenberg & Skaar, 2001). The spelling test battery consists of 79 real words. One point was given for correct response, 0 for incorrect response. The norm score for 6th grade is 55 (SD 15).

L1 Sentence dictation was administered using an especially designed version of TextPilot, a computer-based writing programme (IncludeA/S, 2009), registering all writing activity during the writing of a sentence read out twice to the child. The subjects were asked to write five sentences to dictation, each containing 6 to 12 words. Correctly spelled word earned 1 point, with a maximum score of 42 points. The average score of time used for each word was registered as time score, seconds per word. Importantly, only words that were actually typed went into this analysis, not the number of words in the original dictated sentences (Morken & Helland, 2013).

L2 English was assessed using the three subtasks Word Spelling, Word Reading and Word Translation from the computer-based English 2 Dyslexia Test (Kaasa, Sanne, & Helland, 2004). The English 2 Dyslexia Test reflects the official school curriculum of L2 English in Norwegian elementary school. Focus is on verbal communication with a gradual introduction to written language. The three subtests all focus on the same 22 high frequency English words from different word classes.

Maximum score for each of the three subtests is 22 points. The assessment is compact in the sense that the selected 22 words are included in different texts assessing spelling, reading and comprehension (translation). For a more thorough analysis of the high frequency words used in the task, see Helland (2008).

In the spelling task a sentence is read aloud, followed by instructions to write a target word from the sentence in question. One point is given for each correctly spelt word.

The reading task is to read 10 different sentences for recording. Two or three words within each sentence are assessed. These are the same words as used in the spelling subtest, but the sentences are different. One point is given for each word pronounced correctly.

The task in the translation subtest is to translate the 10 sentences used in the reading subtest from English to Norwegian. One point is given for each correct translation.

Literacy sum scores. To reduce data and compensate for possible ceiling effects, the literacy scores were collapsed in the following manner: The top range score within each variable served as a 100% score. Then, for each variable, a percent score for each participant was calculated: $(\text{variable score} / \text{variable top score}) \times 100$. Finally, each participant's mean percent score within each literacy domain (L1 Read, L1 Write and L2 Skill) was calculated, resulting in three mean sum% scores. For the writing variables (number of errors and seconds/word), low scores are top scores. Hence, scoring for these variables was reversed in order to make them comparable.

Neurocognitive Measures

PA was assessed by the Ringerike Material (Lyster, Tingleff, & Tingleff, 2002), which is a Norwegian test for 5–7-year-old children comparable to, i.e. The Phonological Awareness Test 2 (Robertson & Salter, 2007). The subtests Rhyme, Same phoneme (words beginning with the same phoneme) and Phoneme deletion (deletion of the first phoneme) were used. Reliability coefficient (Spearman–Brown corrected) is 0.92, 0.79 and 0.70, respectively, for the three subtests according to the manual. No validity measures are given, but the test was constructed from tests used in the well-known Bornholm project (Lundberg *et al.*, 1988). In this context only the summed raw scores were used. Correct response was scored with 1 point, incorrect with zero points, yielding a maximum score of 31 points. The children were tested at ages 5 and 6.

STM and WM were assessed by the Digit Span task from the Wechsler Intelligence Scale for Children—Third Edition (Wechsler, 1974, 2003). The task was administered and scored according to test instructions. Raw scores from forward recall (STM) and backward recall (WM) were used because standardized scores were only available from 6 years and only as compound scores. The data collection points were at the ages of 5, 7, 8 and 11. No norms are given for separate forward and backward recall.

Visuo-spatial skills (VS) were assessed by the Rey–Osterrieth Complex Figures Test with the Copy (VS/copy) and Recall (VS/recall) subtests (Meyers & Meyers, 1995; Watanabe *et al.*, 2005). The test was administered according to instructions, but without the 'Immediate recall' condition. The subjects were given a blank piece of paper and a pencil, and were first asked to copy the line drawing figure put in

front of them. After 30 min with other activities they were asked to redraw the figure from memory. The authors of the test underline that not only visuo-spatial abilities are involved in the test performance, but also other functions, such as memory, attention, planning, WM and EF. The 'VS/recall' also involves visual LTM. Each drawing is scored for the accurate reproduction and placement of 18 specific design elements, with a maximum score of 36 points for each task. The data collection points were at ages 6, 7 and 11.

RAN. To accommodate age, two different tests were used. The baseline condition of a Stroop paradigm, RAN colour/word (RAN c/w) (Hugdahl, undated version) was used when the children were 5, 6, 7 and 11 years old. In this test, the participants were shown a sheet with $6 \times 8 = 48$ dots in different colours. The task was to name the colours as quickly and accurately as possible and was scored by timing (seconds) how long the child took to name the colour of all the dots. Timing was done using a stop-watch.

In addition, the RAN task from the STAS test battery, RAN alphanumeric (RANa/n) was used (Klinkenberg & Skaar, 2001). This task involves naming of letters and numbers in scrambled order. Scoring was done by timing how many items the child was able to name correctly within a 40 s limit. Because this test requires letter and number knowledge, the data collection points were at ages 8 and 11.

Verbal LTM was assessed by a verbal learning task. The Vaale Test (Andreassen & Øksenholt, 2009) is a standardized Norwegian equivalent to Luria's 10 word test (Christensen, 1985) and the Children's Auditory Verbal Learning Test—2 (CAVLT-2) measuring auditory verbal learning (Talley, 1993). The test contains 12 words within three categories: clothes, vehicles and animals. All words are read out to the child 10 times, and the child repeats as many words as he or she remembers for each word chain. One point was given for correct response, 0 for incorrect response. Thus, 12 points was the maximum score for each trial. In accordance with learning phases reported by the test constructors (Andreassen & Øksenholt, 2009) only the fourth phase reports LTM after 20 min without any repetition of the list. This task was used when the children were 7 and 11 years old.

Vocabulary was assessed using the British Picture Vocabulary Scale II, a receptive vocabulary test for subjects 3 to 16 years old (Dunn, Whetton, & Burley, 1997). The Norwegian version contains 12 sets with 12 tasks each (Lyster, Horn, & Rygvold, 2010). For each question, the test administrator says a word, and the subject responds by selecting the picture (from four options) that best illustrates the word's meaning. The questions sample words that represent a range of areas, as actions, animals, toys and emotions. The testing was done according to test instructions, and because Norwegian standard scores were not yet available at project start, raw scores were used, with 144 as a maximum score. One point was given for correct response, 0 points for incorrect response. Testing was done at ages 5, 6, 7 and 8.

Sentence comprehension. The Test for Reception of Grammar—Version-2 (TROG-2) (Bishop, 1989) is a receptive language test for subjects 4 to 16 of age. The test contains 80 four-choice items in each of the 20 grammatical blocks arranged in order of increasing difficulty. The test leader reads a sentence, and the subject is to point out the picture that correctly depicts the sentence. The three distractors are contrasted with the target in order to give the test leader insight into grammatical or lexical problems of the subject. The test was scored by block, with a maximum score of 20 points. The data collection points were ages 5, 6, 7, 8 and 11.

Procedures

The assessments were carried out individually in the local communities, either in schools or in the offices of the Educational and Psychological Counselling Services (Pedagogisk-psykologisk tjeneste, PPT). The tests were administered by trained professionals from the local PPT, either by speech and language therapists, special needs teachers or psychologists. Scoring was done by the researchers. Principles for scoring were predetermined, and cases of doubt were discussed in team and agreed upon.

In addition, the children received periodical data-based training in accordance with national curriculums during their last months of preschool and of grade 1 (Pre-literacy stage) as well as the last months of grade 2 (Emergent literacy stage). The reasons for including these training sessions were twofold; (1) to give something back to the children as a compensation for the substantial amount of time they put into their participation, and (2) to provide the schools with updated competence as recognition for the work they put down for the project. The ethics committee would not have approved the project without the training periods, and as such this was a matter of going through with the research or not (Helland, Tjus, Hovden, Ofte, & Heimann, 2011).

Background Data and Design

The gender ratio (M/F) in the Dyslexia group was 5/8, and in the Typical group 16/13. One participant in the Dyslexia group and three participants in the Typical group were reported to be lefthanders. Familial dyslexia was reported in ten children in the Typical group and in eight children in the Dyslexia group.

The study is divided into Part 1 and Part 2. Part 1 assessed first the relationship between the literacy scores and the neurocognitive scores. Second, neurocognitive predictors of Literacy skills at age 11 were assessed. Part 2 assessed developmental group differences in the neurocognitive scores.

The alpha level was set to $p < .05$.

As can be seen in Table 1, there were no group differences in the WPPSI scores, but there were significant group differences in the RI-5 score.

Table 2 shows the literacy scores at age 11. The Dyslexia group scored significantly lower than the Typical group on all L1 and L2 literacy tasks.

Part 1

Two-way correlation (Pearson Product-Moment) was used to assess the relationship between the neurocognitive scores at each of the three literacy stages and the age 11 literacy scores.

To assess neurocognitive predictors of the age 11 literacy scores stepwise linear regression was used. The value of such analyses depends on the correlation between the predictor variables. If predictor variables are too closely related, they may not represent unique contributions to the explained variable R^2 . Because earlier research has shown that PA, VS and RAN are unique contributors in literacy (Kibby, Lee, & Dyer, 2014), and other longitudinal studies have shown that neurocognitive benchmarks in dyslexia vary by age (Dandache, Wouters, & Ghesquière, 2014) regression could be performed. The viability of a regression

Table 2. Dyslexia vs typical group, age 11: L1 and L2 literacy scores

Literacy scores, age 11	Dys	SD	Typ	SD	t-Value	df	p	Cohen's <i>d</i>	N Dys	N Typ
L1 Read										
Read nonwords	48.7	6.01	63.9	13.94	-3.753	39	.001	-1.42	13	28
Read words	151.1	23.70	208.1	40.31	-4.717	39	.001	-1.72	13	28
Text read words/min	80.5	26.71	148.2	32.42	-6.558	39	.001	-2.28	13	28
SUM% L1 read	30.5	5.22	47.7	8.28	-6.851		.001	-2.48	13	28
L1 Write										
Word spelling correct	42.5	5.04	53.6	9.50	-3.949	39	.001	-1.46	13	28
Sent dict errors	11.4	3.43	6.5	2.81	4.830	39	.001	1.56	13	28
Sent dict sec/word	5.2	2.49	3.8	1.20	2.518	38	.002	0.75	13	27
SUM% L1 write	40.5	10.92	50.4	8.08	-3.245		.002	-1.03	13	28
L2 English										
Spelling	6.2	3.22	12.7	4.30	-4.801	39	.001	-1.70	13	28
Read words	14.1	3.38	18.6	2.66	-4.696	39	.001	-1.50	13	28
Translation	18.1	2.96	19.8	1.96	-2.158	39	.040	-0.67	13	28
SUM% L2 English	58.6	12.56	78.3	11.30	-5.012	39	.001	-1.65	13	28

Notes. SUM% scores: see method part.

was also supported by a preliminary analysis showing that the correlation between the neurocognitive scores decreased by age.

For each of the three literacy stages, three separate analyses were performed, using the three age 11 literacy domains as criterion variables. Thus, nine separate regression analyses were performed. Predictor variables were the neurocognitive variables that correlated significantly with the L1 and L2 scores at each literacy stage (see Table 3). If less than three significant correlations were found, the variables with the highest *r*-value in the correlations were used. This was the case for L1 Write at the Emergent literacy stage, where 7RANc/w and 8STM were used in addition to RANa/n, and at the Literacy stage, where 11VS/recall was used in addition to 11 Sentence comprehension and 11WM.

RESULTS

Correlations

The two-way correlation between the neurocognitive variables at each of the three literacy stages and the three age 11 literacy scores is shown in Table 3. As can be seen, the number of significant correlations varied by domain and literacy stage, but also, the number of significant correlations decreased by age.

L1 Read scores correlated with verbal tasks from PA, STM and WM in the Pre-literacy stage, with RAN and LTM in the Emergent literacy stage and with RAN, LTM and Sentence comprehension in the Literacy stage. As to the visual tasks there were significant correlations with the two VS tasks in both Pre-literacy and Emergent literacy stages, but with no correlations in the Literacy stage.

L1 Write correlated significantly with the verbal scores STM, RAN and Sentence comprehension in the Pre-literacy stage, with RAN in the Emergent literacy stage and with WM and Sentence comprehension in the Literacy stage. As to the visuo-spatial tasks a significant correlation was seen in the Pre-literacy stage with VS/recall only.

Table 3. Correlations between neurocognitive scores and L1 and L2 literacy scores

Literacy stage/neurocognitive variable	Literacy domain, age 11		
Pre-literacy	L1 Read	L1 Write	L2 English
5 PA	.125	.232	.433**
6 PA	.314*	.295	.488***
5 STM	.434**	.354*	.371*
5 WM	.340*	.213	.247
6 VS/copy	.454**	.200	.242
6 VS/recall	.490***	.329*	.300
5 RANc/w	-.292	-. .438**	-.239
5 Vocabulary	.264	-.035	.332*
5 Sent. comp.	.092	.355*	.353*
Emergent literacy			
8 STM	.292	.241	.536***
8 WM	.250	.097	.188
7 VS/copy	.310*	.235	.264
7 VS/recall	.350*	.180	.150
7 RANc/w	-. .375*	-.295	-.144
8 RANa/n	.493***	.423**	.323*
7 LTM	.395**	.065	.171
8 Vocabulary	.166	.148	.363*
8 Sent. comp.	.314*	.210	.516***
Literacy			
11 STM	.177	.300	.282
11 WM	.120	.360*	.355*
11 VS/copy	.208	.173	-.044
11 VS/recall	.218	.310	.072
11 RANc/w	.053	-.029	.050
11 RANa/n	.359*	.052	.009
11 LTM	.370*	.175	.449**
11 Sent. comp.	.314*	.402**	.371*

Notes. Number in front of test name = age. PA = phonological awareness; STM = short term memory; WM = working memory; VS/copy = visuo-spatial copy; VS/recall = visuo-spatial recall; RANc/w = RAN colour/word; RAN a/n = RAN alphanumeric; LTM = long term memory; Sent. comp. = sentence comprehension.

Bold data indicate significance.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

L2 English correlated with verbal tasks scores only, and not with any of the visual task scores. In the Pre-literacy stage significant correlations were with PA, STM, Vocabulary and Sentence comprehension, in the Emergent literacy stage with STM, RAN, Vocabulary and Sentence comprehension, and in the Literacy stage with WM, LTM and Sentence comprehension.

Regression Analyses

The regression analyses are shown in Table 4.

Predictors of L1 Read

VS/recall emerged as a significant predictor in both the Pre-literacy and Emergent literacy stages. In the Pre-literacy stage 6VS/recall counted for 24.0% of the variance. In Emergent literacy 8RANa/n and 7VS/recall shared a variance of

Table 4. Linear regression analyses (stepwise) measuring the contribution of neurocognitive scores of literacy scores at each stage

Domain	Stage	Predictor	ANOVA	R ²	Counting for variance %
L1 Read			<i>F</i> (df)	<i>p</i>	
	Pre-literacy	6 VS/recall	<i>F</i> (1,39) = 12.316	.001	.240
	Em literacy	8 RANa/n (1)	<i>F</i> (1,38) = 12.520	.001	.248
		7 VS/recall (2)	<i>F</i> (2,37) = 9.359	.001	.336
	Literacy	11 LTM	<i>F</i> (1,37) = 5.869	.020	.137
L1 Write	Pre-literacy	5 RANc/w	<i>F</i> (1,39) = 9.244	.004	.192
	Em literacy	8 RANa/n	<i>F</i> (1,38) = 10.054	.003	.209
	Literacy	11 Sent. comp (1)	<i>F</i> (1,38) = 7.338	.010	.162
		11 WM (2)	<i>F</i> (2,37) = 7.033	.003	.275
L2 English	Pre-literacy	6 PA	<i>F</i> (1,35) = 7.910	.008	.189
	Em literacy	8 STM (1)	<i>F</i> (1,35) = 15.299	.001	.287
		8 RANa/n (2)	<i>F</i> (2,37) = 12.582	.001	.405
	Literacy	11 LTM	<i>F</i> (1,37) = 9.398	.004	.203

Notes. Variables entered in the model were the neurocognitive variables that correlated significantly with the Baseline 2 scores, see Table 4. In L1 Write the two variables closest to significant correlation, 7 RANc/w and 8 STM, were entered for the Emergent literacy stage analyses and 11 VS/recall was entered for the Literacy stage.

33.6%. 8RANa/n counted for the largest variance in the Emergent literacy stage (24.8%), while VS/recall counted for 8.8% of the variance. In the Literacy stage 11LTM counted for 13.7% of the variance.

Predictors of L1 Write

5RANc/w emerged as a significant predictor at the Pre-literacy stage, counting for 19.2% of the variance, while 8RANa/n counted for 20.9% of the variance in the Emergent literacy stage. At the Literacy stage 11 Sentence comprehension counted for 16.2% of the variance, while 11WM counted for 11.3% of the variance.

Predictors of L2 English

In the Pre-literacy stage 6PA counted for 18.9% of the variance. In the Emergent literacy stage 8STM counted for 28.7% of the variance and 8RANa/n for 11.8% of the variance. In the Literacy stage 11LTM counted for 20.3% of the variance.

Summary: Part I, Correlations and Predictors

All the neurocognitive variables correlated significantly with the age 11 literacy scores, but the significance differed by literacy domain (L1 Read, L1 Write and L2 English), and by literacy stages. This tendency was also seen in the regression analyses. In the Pre-literacy stage the emerging predictors were VS/recall in L1 Read and RAN in L1 Write, and PA in the L2 English. In the Emergent literacy stage RAN emerged as the significant predictor of all three domains, with an addition of VS/recall in L1 Read and STM in L2 English. In the Literacy stage LTM was a

significant predictor of L1 Read and L2 English. Sentence comprehension and WM were significant predictors of L1 Write.

Part 2

Longitudinal neurocognitive scores by group

For group comparisons and longitudinal analyses the neurocognitive tests were subjected to repeated measures ANOVA with the design test scores by Group (Dyslexia, Typical) by Age. Not all tests were used at each assessment point. This was due to the total amount of tests for each child (they were also tested with other tasks not reported here), and due to the age standardization of the tests. Significant effects from ANOVA were followed up by Fisher's LSD test. In addition, two-tailed t-tests (Typical vs Dyslexia) were executed for each test point.

Results

Repeated measures ANOVA and t-tests

Repeated measures for all neurocognitive tasks are shown in Figure 1. For all repeated measures analyses there were significant effects of age ($p < .0001$). Effects of Group ($p < .05$) were seen in STM, WM, VS/copy, RAN c/w and RAN a/n. Tendencies (effects $p = .06 - .10$) were seen in PA, VS/recall, LTM and Sentence comprehension, but not in Vocabulary.

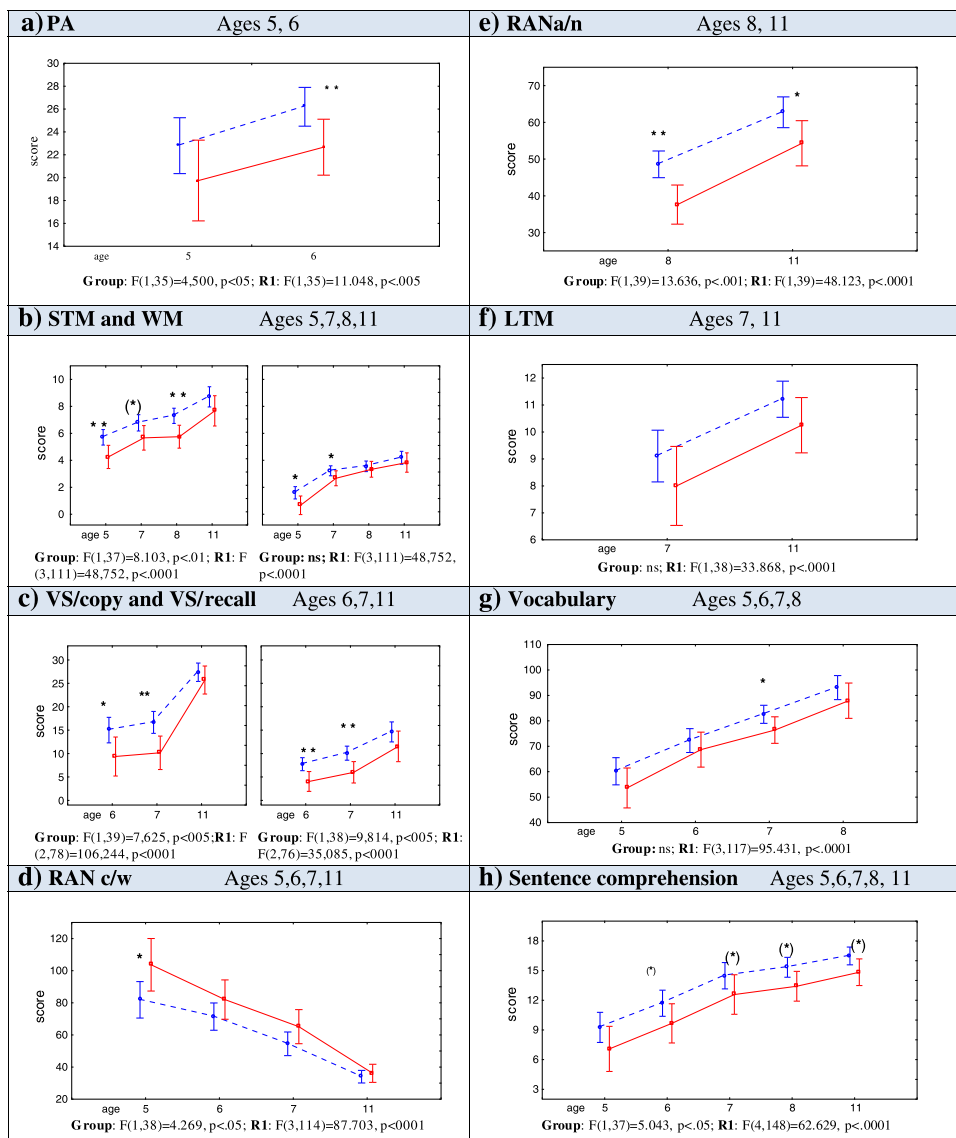
Cross-sectional t-test (two-tailed) showed that the effects of Group were mainly because of significant differences in the Pre-literacy and Emergent literacy stages, i.e. between ages 5 and 8. At age 11 (Literacy stage) these significant differences disappeared in all tasks, except for in RANa/n and a close to significance in Sentence comprehension. However, when using one-tailed t-test significance was reached ($p < .05$), and Cohen's d ranged in the four age tests from .54 to 0.69. Similarly, in STM, age 7, when using one-tailed t-test significance was reached ($p < .05$), and Cohen's d was 0.74, which is defined to be in the area medium to large.

Summary: Part 2, Neurocognitive Development by Age

The neurocognitive scores improved by age in both groups, and the scores of the Dyslexia group were in general significantly lower than the Typical group. The group differences seen in the Pre-literacy and Emergent literacy stages declined in the Literacy stage, and were only seen in RANa/n.

DISCUSSION

In the present study we first wanted to assess the correspondence between benchmark neurocognitive skills in dyslexia across the Pre-literacy, Emergent literacy and Literacy stages and literacy skills defined by the three domains L1 reading, L1 writing and L2 English at age 11. Second, at each of the three stages of age 11 literacy skills we wanted to find valid neurocognitive precursors. Third, we wanted to compare the neurocognitive development in the Dyslexia group and the Typical group across the three stages of L1 and L2 literacy acquisition. In sum a complex



Notes. Significant differences (two-tailed t-test) marked with *Dys < Typ*: $p<.05$; **: $p<.01$; ***: $p<.001$.

p<.001. Significant differences (one-tailed t-test) marked with (*).Dys < Typ*: $p<.05$ (*) where Cohen's d on Sentence comprehension ranged in the four age tests from .54 to 0.69 and on STM, age 7, Cohen's d was 0.74. Abbreviations as in Table 3.

Figure 1. Repeated measures ANOVA by group. Solid line: Dyslexia group, dotted line: Typical group. Notes. Significant differences (two-tailed t-test) marked with * Dys < Typ*: $p<.05$; **: $p<.01$; ***: $p<.001$. Significant differences (one-tailed t-test) marked with (*). Dys < Typ*: $p<.05$ (*) where Cohen's d on Sentence comprehension ranged in the four age tests from .54 to 0.69 and on STM, age 7, Cohen's d was 0.74. Abbreviations as in Table 3.

developmental picture emerged, which may explain how outcome differences in dyslexia research could be understood. This complex picture is first tentatively illustrated by a model as shown in Figure 2.

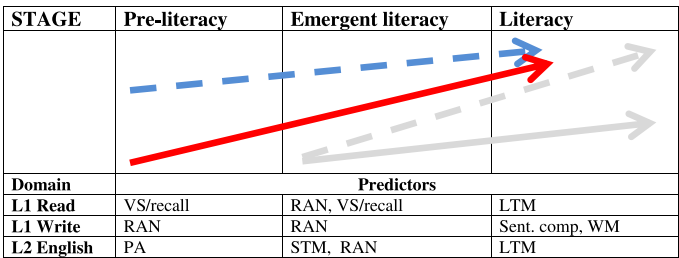


Figure 2. A model of neurocognitive and literacy development and significant predictors at each literacy stage. Greylines illustrate collapsed literacy scores; coloured lines illustrate collapsed neurocognitive scores. Solid lines: Dyslexia group; dotted lines: Typical group. Abbreviations as in Table 4.

The coloured lines illustrate how the neurocognitive scores developed from showing a significant gap between the groups in the Pre-literacy and Emergent literacy stages to displaying only minor differences between groups in the Literacy stage. Parallel to this development the grey lines illustrate how the Dyslexia group gradually fell behind the Typical group as to literacy skills, even if they had gained functional, but not very effective, literacy skills at this age. The development is discussed by literacy stage.

Pre-Literacy and Emergent Literacy Stages

VS/recall in L1 Read

In the Pre-literacy stage visuo-spatial memory emerged as the most prominent predictor of reading skills at age 11. This was surprising seeing that the role of the VS in dyslexia has been controversial; it has been acknowledged (Boder, 1973), debated (Vellutino, 1978), rejected (Vellutino *et al.*, 2004), and revitalized (Stein & Walsh, 1997; Valdois *et al.*, 2012). Most studies have not considered VS as essential to pre-school language development or language impairment. However, that the VS/recall variable in the Pre-literacy stage emerged as especially important for later L1 reading skills is in line with the earlier reported tendency of younger children to remember pictorial information in visual and not verbal formats (Gathercole & Baddeley, 1993), and the findings of Huestegge *et al.* who found a deficit in visual LTM in older dyslexic children (Huestegge, Rohrßen, von Ermingen-Marbach, Pape-Neumann, & Heim, 2014).

This finding is also in line with our retrospective analyses of cortical thickness and neuroanatomical precursors in the Dyslexia and Typical groups at age 6, indicating that dyslexia could be caused by structural deficiencies in the lower-level networks responsible for auditory and visual processing and core EF, and not from impaired literacy functions (Clark *et al.*, 2014). In sum this shows that more attention should be given to VS in the Pre-literacy stage both in research and clinical work.

RAN in L1 Write, L1 Read and L2 English

The early predictive value of the RAN scores as seen here is in line with much dyslexia research, however mainly reported in impaired reading (Denckla & Cutting, 1999; Kirby *et al.*, 2010; Norton & Wolf, 2012). In both the Pre-literacy and Emergent literacy stages RAN turned out to be the sole significant predictor of

L1 Write, but also of L1 Read and L2 English in the Emergent literacy stage. This is in line with other reports on writing in both children and adults with dyslexia (Berninger, Nielsen, Abbott, Wijsman, & Raskind, 2008) and with reports from other orthographies (Yeung, Ho, Chan, & Chung, 2014). Interestingly, Amtmann *et al.* found a relationship between spelling and RAN, also indicating that RAN training and spelling should go hand in hand (Amtmann, Abbott, & Berninger, 2008).

PA and STM in L2

PA has for long been seen as the core at-risk factor of dyslexia. Phonology is the study of sound systems of individual languages, and of the nature of such systems generally (Matthews, 1997). A striking result in the present study was PA as a predictor at the Pre-literacy stage of L2 English, and not of L1 Norwegian. In the Emergent literacy stage STM appeared as the best L2 predictor. PA was not tested at this stage, but in the literature STM is often seen as under the phonological skills umbrella (Melby-Lervåg, Lyster, & Hulme, 2012), or as a cause of phonological skills (Tønnessen & Uppstad, 2015; Uppstad & Tønnessen, 2007). Learning a new language puts demands on the perception and production of unfamiliar phonemes. English contains several phonemes that are unfamiliar in Norwegian, such as /θ/ as in 'think', /ð/ as in 'this', /w/ as in 'wind', /z/ as in 'zoo', /dʒ/ as in 'just' and /ʒ/ as in 'genre'. This difference may be one way of explaining the impact of PA and STM in English as L2. The impact of impaired PA and STM on writing in a highly irregular orthography like English cannot be evaluated at this stage, because the focus in schools is on oral L2 language according to school curriculum. Still, one may infer that low PA and STM scores in the Pre-literacy and Emergent literacy stages have major consequences for later L2 English reading and writing competence. Also, this is an indication that PA in reading and writing acquisition is sensitive to orthography, rather than a universal benchmark of dyslexia. A bit surprising was the lack of VS as a predictor of a deep orthography as L2 English. However, as described in the introductory part, the focus on oral communication and pragmatics in the elementary school may explain this finding. This picture may change in high school, when reading and writing in English become main parts of the curriculum.

Literacy Stage

LTM, sentence comprehension and WM

In the present study LTM emerged as a predictor of both L1 Read and L2 English in the Literacy stage. As described earlier, at this stage reading and writing are expected to be tools for academic learning, and not skills to be learned in themselves any more. Both text reading and second language learning build on linguistic recognition and memory, defined as components of the episodic buffer (Baddeley, 2000). Likewise, learning the highly irregular correspondence between phonemes and graphemes of the English orthography requires a high number of repetitions to be stored in LTM. Usually, children with dyslexia read and write less than typical children, and thus their statistical learning process is slower both as to vocabulary and spelling acquisition.

Sentence comprehension and WM emerged as predictors of L1 writing skills at the Literacy stage. The writing score was based on single word and sentence dictation, and the scoring included both a product (spelling and semantic errors)

and a process score (transcription fluency and revisions) (Morken & Helland, 2013). The writing of free narratives was not assessed here. Because the longitudinal group differences were minor in these three neurocognitive variables (WM, LTM and Sentence comprehension), it would be of interest to assess these variables in more demanding literacy tasks like text reading and compositional writing.

In a previous fMRI-study of a subgroup of these children (Morken *et al.*, 2014), we investigated the effect of increasing literacy demands upon reading-related neuronal activation. We found that the dyslexia group responded with relatively increased activity in areas that are associated with sentence comprehension and processing, as well as higher cognitive functions such as for example attention and language monitoring, rather than in the reading network (see e.g. Dehaene, 2009; Sandak, Mencl, Frost, & Pugh, 2004) *per se*. We interpreted this as reliance upon compensatory mechanisms in the dyslexia group, increasing with task complexity. This held for both specifically linguistic functions and for more general neurocognitive functions important for language processing. These findings support the need for more research into the effect of neurocognitive skills on performance in higher-level literacy tasks.

CONCLUDING REMARKS

It may seem contradictory that the group differences in the literacy scores increased by literacy stage, while the group differences in the neurocognitive scores decreased by literacy stage. According to earlier research, learning to read changes brain responses in the visual occipital areas and the language network in the left hemisphere (see i.e. Dehaene *et al.* (2010) and Carreiras *et al.* (2009)). Analogous to this, adult illiterates are shown to have improved their phonological skills after having learned to read and write (Morais, 1991). It is reasonable to assume that literacy training in itself interacted with and improved the children's neurocognitive skills. Hence, at age 11 the relatively low literacy scores in the Dyslexia group could be explained by the early impairments in visuo-spatial memory, RAN (in L1), and PA and STM (in L2) in the Pre-literacy and Emergent literacy stages. The present findings could therefore underline the importance of detecting early at-risk signs followed by evidence-based training.

Initially it was pointed out that dyslexia is characterized by a number of neurocognitive deficits. Leaning on well accepted neurocognitive benchmarks in dyslexia the longitudinal approach in the present study shed new light on how these benchmarks change and interact with developing literacy skills. Our results support many of the earlier findings reported in the introduction, but the longitudinal analyses also brought some unexpected results. One reason could be found in the longitudinal design. As proposed by Goswami, our developmental design has been within a neuroconstructivist framework, and a low-level impairment has been identified at a Pre-literacy stage, giving a basis for the understanding of the development of higher-level cognition (Goswami, 2003, 2008). Another reason may be that dyslexia most often is defined by reading impairment, while it in the present study was defined by both L1 and L2 reading and writing competence. The novelty was that each of these domains exhibited different neurocognitive developmental patterns.

The longitudinal approach also showed a neurocognitive development which can be framed within the Model of Working Memory (Baddeley, 2000). In the Pre-literacy and the Emergent literacy stages impairments were seen within the phonological loop and visuo-spatial sketchpad, and in the Literacy stage impairments were seen within the episodic buffer. A biological framework was first that our neurocognitive findings corresponded with our retrospective analyses of cortical thickness at age 6, indicating that dyslexia most probably is caused by structural deficiencies in the lower-level networks responsible for auditory and visual processing and core EF, and not from impaired literacy functions (Clark *et al.*, 2014). Second, the cortical overactivation seen in the Dyslexia group at age 11 (Morken *et al.*, 2014) support reliance upon linguistic and non-linguistic neurocognitive functions in more complex literacy tasks, even in the Literacy stage when an acceptable level of proficiency has been reached. Also, seen within this framework the precursors went from lower-level to higher-level cognitive functions.

Both the correlative and comparative group designs showed normalization of neurocognitive functions by age. This implies that the impaired neurocognitive scores in the Pre- and Emergent literacy stages should be seen as early endophenotype precursors of the development of dyslexia. The fact that both groups received training in the early phases of the project (Helland *et al.*, 2011) may seem confounding from a scientific point of view. However, if a demanding research project involving children is to be implemented in a pedagogical setting, it must have a face value to the participants. This means that the research project has to be meaningful to the involved schools, teachers, children and parents. Even though the training may have facilitated and prepared the brain's networks for learning to read and write, the literacy differences between the groups remained highly significant. The present findings should therefore underline the importance of detecting early at-risk signs followed by evidence based training.

Thus, further studies and clinical assessments should most of all focus on these early, lower-level neurocognitive functions, both for early detection and for early intervention in children at-risk of developmental dyslexia.

Practitioner Points:

- Pre-literacy, Emergent literacy and Literacy stages showed different neurocognitive patterns in accordance with domain (L1 read, L write, L2 English)
- In the Pre-literacy stage visuo-spatial skills predicted L1 reading, RAN predicted L1 writing and PA predicted L2 English
- In the Emergent stage RAN predicted all three domains. Also, visuo-spatial skills predicted L1 reading and STM predicted L2 English
- In the Literacy stage LTM predicted L1 reading and L2 English, while Sentence comprehension and WM predicted L1 writing
- This development from lower to higher level neurocognitive precursors should be accounted for in dyslexia assessments

ACKNOWLEDGEMENTS

The project was supported by grants from the University of Bergen, Statped Vest, the Norwegian Research Council and the municipalities Haugesund, Kvinnherad, Førde and Fræna.

REFERENCES

- Amtmann, D., Abbott, R. D., & Berninger, V. W. (2008). Identifying and predicting classes of response to explicit phonological spelling instruction during independent composing. *Journal of Learning Disabilities*, 41(3), 218–234. doi:10.1177/0022219408315639.
- Andreassen, T. H., & Øksenholt, S. I. (2009). Våletesten [the vaale test]. *Keisertveien* 35, 3135 Torud: (Available from: <http://www.våletesten.no>).
- APA (1994). *DSM-IV. Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association.
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417–423.
- Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, 36(3), 189–208. doi:10.1016/S0021-9924(03)00019-4.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.
- Beneventi, H., Tønnesen, F. E., & Ersland, L. (2009). Dyslexic children show short-term memory deficits in phonological storage and serial rehearsal: An fMRI study. *International Journal of Neuroscience*, 119, 2017–2043. doi:10.1080/00207450903139671.
- Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Writing problems in developmental dyslexia: Under-recognized and under-treated. *Journal of School Psychology*, 46(1), 1–21. doi:10.1016/j.jsp.2006.11.008.
- Bishop, D. V. M. (1989). Test for reception of grammar. In 2008 Norwegian translation by Solveig-Alma Lyster (2nd ed.). Manchester: University of Manchester: The Author, Age and Cognitive Research Centre.
- Bishop, D. V. M., & Leonard, L. B. (2000). Speech and language impairments in children. Causes, characteristics, intervention and outcome. Hove: Psychology Press.
- Blomert, L., & Willems, G. (2010). Is there a causal link from a phonological awareness deficit to reading failure in children at familial risk for dyslexia? *Dyslexia*, 16(4), 300–317. doi:10.1002/dys.405.
- Boder, E. (1973). Developmental dyslexia: A diagnostic approach based on three atypical reading–spelling patterns. *Developmental Medicine & Child Neurology*, 15(5), 663–687. doi:10.1111/j.1469-8749.1973.tb05180.x.
- Carlsten, C. T. (2002). *Leseprøve 6. klasse bokmål og nynorsk [Reading test 6th grade bokmål and nynorsk]*. Oslo: Damm & Søn AS.
- Carreiras, M., Seghier, M. L., Baquero, S., Estevez, A., Lozano, A., Devlin, J. T., & Price, C. J. (2009). An anatomical signature for literacy. *Nature*, 461(7266), 983–986. DOI:10.1038/nature08461. (Available from: http://www.nature.com/nature/journal/v461/n7266/supinfo/nature08461_S1.html)
- Christensen, A.-L. (1985). *Luria's neuropsychological investigation*. Copenhagen: Munksgaard.
- Chung, K. K. H., & Ho, C. S.-H. (2010). Second language learning difficulties in Chinese children with dyslexia: what are the reading-related cognitive skills that contribute to English and Chinese word reading? *Journal of Learning Disabilities*, 43(3), 195–211.
- Clark, K. A., Helland, T., Specht, K., Narr, K. L., Manis, F. R., Toga, A. W., & Hugdahl, K. (2014). Neuroanatomical precursors of dyslexia identified from pre-reading through to age 11. *Brain*. doi:10.1093/brain/awu229.
- Dahl, A., & Vulchanova, M. D. (2014). Naturalistic acquisition in an early language classroom. *Frontiers in Psychology*, 5, 329. doi:10.3389/fpsyg.2014.00329.
- Dandache, S., Wouters, J., & Ghesquière, P. (2014). Development of reading and phonological skills of children at family risk for dyslexia: A longitudinal analysis from kindergarten to sixth grade. *Dyslexia*. doi:10.1002/dys.1482.
- Dehaene, S. (2009). *Reading in the brain*. New York: Viking.
- Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., ... Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, 330(6009), 1359–1364. doi:10.1126/science.1194140.

- Denckla, M., & Cutting, L. (1999). History and significance of rapid automatized naming. *Annals of Dyslexia*, 49(1), 29–42. doi:10.1007/s11881-999-0018-9.
- Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British picture vocabulary scale*. Second Edition. London, UK: Nelson Publishing Company.
- Ehri, L. C. (1987). Learning to read and spell words. *Journal of Literacy Research*, 19(1), 5–31. doi:10.1080/10862968709547585.
- Frith, U. (1986). A developmental framework for developmental dyslexia. *Annals of Dyslexia*, 36, 69–81.
- Frith, U. (1995). Dyslexia: Can we have a shared theoretical framework? *Educational & Child Psychology*, 12(1), 6–17.
- Frith, U. (1999). Paradoxes in the definition of dyslexia. *Dyslexia*, 5, 192–214.
- Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A.-M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology*, 93(3), 265–281. doi:10.1016/j.jecp.2005.08.003.
- Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. East Sussex: Psychology Press Ltd.
- Geva, E., Yaghouz Zadeh, Z., & Schuster, B. (2000). Understanding individual differences in word recognition skills of ESL children. *Annals of Dyslexia*, 50, 123–154.
- Gooch, D., Snowling, M., & Hulme, C. (2011). Time perception, phonological skills and executive function in children with dyslexia and/or ADHD symptoms. *Journal of Child Psychology and Psychiatry*, 52(2), 195–203. doi:10.1111/j.1469-7610.2010.02312.x.
- Goswami, U. (2003). Why theories about developmental dyslexia require developmental designs. *Trends in Cognitive Sciences*, 7(12), 534–540. doi:10.1016/j.tics.2003.10.003.
- Goswami, U. (2008). Reading, dyslexia and the brain. *Educational Research*, 50(2), 135–148. doi:10.1080/00131880802082625.
- Halliday, L. F. (2014). A tale of two studies on auditory training in children: A response to the claim that 'discrimination training of phonemic contrasts enhances phonological processing in mainstream school children' by Moore, Rosenberg and Coleman (2005). *Dyslexia*, 20(2), 101–118. doi:10.1002/dys.1470.
- Halvorsen, K., & Stjernø, S. (2008). *Work, oil and welfare*. Oslo: Universitetsforlaget.
- Heiervang, E., & Hugdahl, K. (2003). Impaired visual attention in children with dyslexia. *Journal of Learning Disabilities*, 36(1), 68–73.
- Helland, T. (2007). Dyslexia at a behavioural and a cognitive level. *Dyslexia*, 13(1), 25–41. doi:10.1002/dys.325.
- Helland, T. (2008). Second language assessment in dyslexia: Principles and practice Language Learners with Special Needs. An International Perspective (pp. 63–85): Multilingual Matters.
- Helland, T. (2015). RI-5. *Dyslexia Risk Index*. Bryne, Norway: InfoVest.
- Helland, T., & Asbjørnsen, A. E. (2000). Executive functions in dyslexia. *Child Neuropsychology*, 6(1), 37–48.
- Helland, T., & Asbjørnsen, A. E. (2003). Visual-sequential and visuo-spatial skills in dyslexia: Variations according to language comprehension and mathematics skills. *Child Neuropsychology*, 9(3), 208–220.
- Helland, T., & Asbjørnsen, A. E. (2004). Digit span in dyslexia: Variations according to language comprehension and mathematics skills. *Journal of Clinical and Experimental Neuropsychology*, 26(1), 31–42.
- Helland, T., & Kaasa, R. (2005). Dyslexia in English as a second language. *Dyslexia*, 11(1), 41–60. doi:10.1002/dys.286.
- Helland, T., Plante, E., & Hugdahl, K. (2011). Predicting dyslexia at age 11 from a risk index questionnaire at age 5. *Dyslexia*, 17(3), 207–226. doi:10.1002/dys.432.
- Helland, T., Tjus, T., Hovden, M., Ofte, S., & Heimann, M. (2011). Effects of bottom-up and top-down intervention principles in emergent literacy in children at risk of developmental dyslexia: A longitudinal study. *Journal of Learning Disabilities*, 44(2), 105–122. doi:10.1177/0022219410391188.
- Huestegge, L., Rohrßen, J., von Ermingen-Marbach, M., Pape-Neumann, J., & Heim, S. (2014). Devil in the details? Developmental dyslexia and visual long-term memory for details. [Original Research]. *Frontiers in Psychology*, 5. doi:10.3389/fpsyg.2014.00686.

Include A/S. (2009). TextPilot. Bergen: Include as.

Kaasa, R., Sanne, S., & Helland, T. (2004). The English 2 dyslexia test. Bergen, Norway: (Available from: <http://www.vesttest.no>).

Kibby, M. Y., & Cohen, M. L. (2008). Memory functioning in children with reading disabilities and/or attention deficit/hyperactivity disorder: A clinical investigation of their working memory and long-term functioning. *Child Neuropsychology*, 14, 525–546.

Kibby, M. Y., Lee, S. E., & Dyer, S. M. (2014). Reading performance is predicted by more than phonological processing. [Original Research]. *Frontiers in Psychology*, 5. doi:10.3389/fpsyg.2014.00960.

Kirby, J. R., Georgiou, G. K., Martinussen, R., & Parrila, R. (2010). Naming speed and reading: From prediction to instruction. *Reading Research Quarterly*, 45(3), 341–362. doi:10.1598/RRQ.45.3.4.

Klinkenberg, J. E., & Skaar, E. (2001). STAS. Standardisert test i avkodning og staving [Standardised test in decoding and spelling]. Hønefoss: Pedagogisk-psykologisk tjeneste.

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., . . . Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry*, 54(6), 686–694. doi:10.1111/jcpp.12029.

Landerl, K., & Willburger, E. (2010). Temporal processing, attention, and learning disorders. *Learning and Individual Differences*, 20(5), 393–401. doi:10.1016/j.lindif.2010.03.008.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, 53(1), 1–14. doi:10.1007/s11881-003-0001-9.

Lyster, S.-A. H., Horn, E., & Rygvold, A.-L. (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge. Resultater fra en utprøving av British picture vocabulary scale, second edition (BPVS II). . *Spesialpedagogikk*, 9, 35–43.

Lyster, S. A. H., Tingleff, H., & Tingleff, Ø. (2002). *Ringeriksmaterialet [The Ringerik Test]*. Oslo: Damm.

Matthews, P. H. (1997). *The concise Oxford dictionary of linguistics*. Oxford: Oxford University Press.

Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading*, 34(1), 114–135. doi:10.1111/j.1467-9817.2010.01477.x.

Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, 138(2), 322–352. doi:10.1037/a0026744.

Meyers, J. E., & Meyers, K., R. (1995). *Rey complex figure test and recognition trial*. Odessa, FL: Psychological Assessment Resources, Inc.

Miller-Guron, L., & Lundberg, I. (2000). Dyslexia and second language reading: A second bite at the apple? *Reading and Writing*, 12(1–2), 41–61. doi:10.1023/A:1008009703641.

Moll, K., Loff, A., & Snowling, M. J. (2013). Cognitive endophenotypes of dyslexia. *Scientific Studies of Reading*, 17(6), 385–397. doi:10.1080/10888438.2012.736439.

Morais, J. (1991). Phonological Awareness: A bridge between language and literacy. In D. J. Sawyer & B. J. Fox (Eds.), *Phonological awareness in reading. The evolution of current perspective* (pp. 31–71). New York: Springer-Verlag.

Morken, F., & Helland, T. (2013). Writing in dyslexia: Product and process. *Dyslexia*, 19(3), 131–148. doi:10.1002/dys.1455.

Morken, F., Helland, T., Hugdahl, K., & Specht, K. (2014). Children with dyslexia show cortical hyperactivation in response to increasing literacy processing demands. *Frontiers in Psychology*, 5. doi:10.3389/fpsyg.2014.01491.

Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology*, 63(1), 427–452. doi:10.1146/annurev-psych-120710-100431.

Norwegian Directorate for Education and Training. (2005). *Kindergarten act*. Oslo: The Norwegian Directorate for Education and Training. (Available from <http://www.regjeringen.no/en/dokumenter/kindergarten-act/id115281/>)

Norwegian Directorate for Education and Training. (2009). *Pisa 2009*. Oslo. (Available from <http://www.udir.no/Tilstand/Internasjonale-studier-/PISA-2009-Bedre-resultater-for-norske-elever-/>)

- Norwegian Directorate for Education and Training. (2013). *English subject curriculum*. Oslo (Available from <http://www.udir.no/kl06/ENGI-03>)
- Norwegian Directorate for Education and Training. (2010–2011). *Learning together*. Oslo. (Available from <https://www.regjeringen.no/nb/dokumenter/meld-st-18-20102011/id639487/>)
- Pancsofar, N., & Vernon-Feagans, L. (2006). Mother and father input to young children: Contribution to later language development. *Journal of Applied Developmental Psychology*, 27, 571–587.
- Pennington, B. F., & Bishop, D. V. M. (2009). Relations among speech, language, and reading disorders. *Annual Review of Psychology*, 60(1), 283–306. doi:10.1146/annurev.psych.60.110707.163548.
- Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology*, 13(2), 212–218. doi:10.1016/S0959-4388(03)00035-7.
- Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *The Quarterly Journal of Experimental Psychology*, 61(1), 129–141. doi:10.1080/17470210701508822.
- Reiter, A., Tucha, O., & Lange, K. W. (2005). Executive functions in children with dyslexia. *Dyslexia*, 11(2), 116–131. doi:10.1002/dys.289.
- Robertson, C., & Salter, W. (2007). *The phonological awareness test 2*. East Moline, IL: LinguSystems.
- Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., Francis, D. J., Pugh, K. R. & Shaywitz, B. A. (1999). Persistence of dyslexia: The Connecticut longitudinal study at adolescence. *Pediatrics*, 104(6), 1351–1359. doi:10.1542/peds.104.6.1351.
- Schwartz, R. G. (Ed.). (2009). *Handbook of child language disorders*. New York and Hove: Psychology Press.
- Snowling, M., Bishop, D. V. M., & Stothard, S. E. (2000). Is preschool language impairment a risk factor for dyslexia in adolescence? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 41(5), 587–600.
- Specht, K., Hugdahl, K., Ofte, S., Nygaard, M., Bjoernerud, A., Plante, E., & Helland, T. (2009). Brain activation on pre-reading tasks reveals at-risk status for dyslexia in 6-year-old children. *Scandinavian Journal of Psychology*, 50(1), 79–91. doi:10.1111/j.1467-9450.2008.00688.x.
- Stein, J., & Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends in Neuroscience*, 20(4), 147–152.
- Talley, J. L. (1993). *Children's auditory verbal learning test—2*. USA: Psychological Assessment Resources, Inc..
- The British Dyslexia Association, B. (2007). British dyslexia association: Definition of dyslexia, (Available from: <http://www.bdadyslexia.org.uk/about-dyslexia/faqs.html>). Retrieved 08.12.08, 2008.
- Torgesen, J. K., & Houck, D. G. (1980). Processing deficiencies of learning-disabled children who perform poorly on the digit span test. *Journal of Educational Psychology*, 72, 141–160.
- Torppa, M., Lyytinen, P., Erskine, J., Eklund, K., & Lyytinen, H. (2010). Language development, literacy skills, and predictive connections to reading in Finnish children with and without familial risk for dyslexia. *Journal of Learning Disabilities*, 43(4), 308–321. doi:10.1177/0022219410369096.
- Trecy, M. P., Steve, M., & Martine, P. (2013). Impaired short-term memory for order in adults with dyslexia. *Research in Developmental Disabilities*, 34(7), 2211–2223. doi:10.1016/j.ridd.2013.04.005.
- Tønnessen, F. E. (1997). How can we best define 'dyslexia'? *Dyslexia*, 3(2), 78–92. doi:10.1002/(SICI)1099-0909(199706)3:2<78::AID-DYS71>3.0.CO;2-2.
- Tønnessen, F. E., & Uppstad, P. H. (2015). *Can we read letters?* Reflections on fundamental issues in reading and dyslexia research.
- Uppstad, P. H., & Tønnessen, F. E. (2007). The notion of 'phonology' in dyslexia research: Cognitivism—and beyond. *Dyslexia*, 13(3), 154–174. doi:10.1002/dys.332.
- Valdois, S., Lassus-Sangosse, D., & Lobier, M. (2012). Impaired letter-string processing in developmental dyslexia: What visual-to-phonology code mapping disorder? *Dyslexia*, 18(2), 77–93. doi:10.1002/dys.1437.
- Vellutino, F. R. (1978). *Visual processing deficiencies in poor readers: A critique of traditional conceptualizations of the etiology of dyslexia*.
- Vellutino, F. R., Fletcher, L. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, 45(1), 2–40.

- Vidyasagar, T. R., & Pammer, K. (2010). Dyslexia: A deficit in visuo-spatial attention, not in phonological processing. *Trends in Cognitive Sciences*, 14(2), 57–63. doi:10.1016/j.tics.2009.12.003.
- Vulchanova, M., Foy, C. H., Nilsen, R. A., & Sigmundsson, H. (2014). Links between phonological memory, first language competence and second language competence in 10-year-old children. *Learning and Individual Differences*(0). doi:10.1016/j.lindif.2014.07.016.
- Wang, Z., Cheng-Lai, A., Song, Y., Cutting, L., Jiang, Y., Lin, O., . . . Zhou, X. (2014). A perceptual learning deficit in Chinese developmental dyslexia as revealed by visual texture discrimination training. *Dyslexia*, 20(3), 280–296. doi:10.1002/dys.1475.
- Watanabe, K., Ogino, T., Nakano, K., Hattori, J., Kado, Y., Sanada, S., & Ohtsuka, Y. (2005). The Rey–Osterrieth Complex Figure as a measure of executive function in childhood. *Brain and Development*, 27(8), 564–569. doi:10.1016/j.braindev.2005.02.007.
- Wechsler, D. (1974). Wechsler intelligence scale for children—revised (Norwegian Edition by Undheim, J. O. ed.): Jaren, Norway: Vigga-trykk.
- Wechsler, D. (2002). Wechsler preschool and primary scale of intelligence™—third edition (WPPSI™ – III). In *Norsk tilpasning*. San Antonio, Texas: Pearson.
- Wechsler, D. (2003). *WISC-III. Norwegian version*. Stockholm: Assessio Norge AS.
- Xiao, X.-Y., & Ho, C. S.-H. (2013). Weaknesses in semantic, syntactic and oral language expression contribute to reading difficulties in Chinese dyslexic children. *Dyslexia*. doi:10.1002/dys.1460.
- Yeung, P.-s., Ho, C. S.-h., Chan, D. W.-o., & Chung, K. K.-h. (2014). What are the early indicators of persistent word reading difficulties among Chinese readers in elementary grades? *Dyslexia*. doi:10.1002/dys.1471.